5. THE APPLICATION OF HUMAN MODELING TECHNOLOGY TO THE DESIGN, EVALUATION AND OPERATION OF COMPLEX SYSTEMS[1]

Wayne Zachary, Gwendolyn E. Campbell,
K. Ronald Laughery, Floyd Glenn and
Janis A. Cannon-Bowers

ABSTRACT

This chapter reviews the ability of the emerging human performance modeling technologies to support the design and operation of complex systems. The ability of existing technologies to meet current application needs is analyzed, and the results are then used to assess the areas where additional research and development is most needed. Following a brief history of human performance modeling, a taxonomy of models and modeling techniques is established, as a framework for remaining discussion. The human performance modeling technology base is separately analyzed for its ability to support system design processing and to support system operation. The system design process analysis considers the various roles that human performance models may play during that

Advances in Human Performance and Cognitive Engineering Research, Volume 1, pages 199–247.

ISBN: 0-7623-0748-X

process, ranging from generating design concepts to affording simulation-based design evaluation. The system operation analysis also assesses a range of roles, from training to performance support to automation. These analyses demonstrate that human modeling technology has reached a sufficient state of maturity that has become a proven contributor of the complex systems engineering process. Challenges for further high-payoff research are also presented, in five categories: cognition, knowledge management, team and organizational structure and processes, predictive models of training, and human-centered systems engineering.

INTRODUCTION

A continuing activity in behavioral and social science research has been constructing models as a means of formalizing, integrating, testing, and even developing theory. In recent years, this avenue of research has been very active, with substantial advances in the development of computer models of human thought and behavior (see Pew & Mavor, 1998). This research has begun to fuel speculation that the human modeling endeavor may have progressed to the point that the current generation of models could have immediate or near-term engineering application.

At the same time, both industry and the government (particularly the military) are facing challenges that are driving the need for human performance models. For example, DoD is requiring that future Navy ships must be built with reduced budgets *and* operated by vastly smaller crews, and must operate effectively and efficiently in mission environments that are complex, difficult to define, and rapidly changing. Thus, the human component is, more than ever, the critical component to mission success. Moreover, the sheer size and scale of complex systems such as ships (or factories, or aircraft, etc.) severely limits the opportunity for traditional experimental approaches. It is simply too costly and time consuming to build physical prototypes (or even effective mock-ups) and then empirically assess their impact on human performance. In this context, the techniques of human modeling and simulation provide perhaps the only viable option. They hold the promise of providing powerful new means to design, evaluate and operate modern ship systems to meet, accommodate, and enhance human abilities, if the technology exists to model them appropriately and effectively. Recent successes, such as the use of human performance models in embedded Naval training (see Zachary, Bilazarian, Burns & Cannon-Bowers, 1997; Zachary, Cannon-Bowers, Burns, Bilazarian & Krecker, 1998) and in support of large scale training exercises (see Laird, Coulter, Jones, Kenny, Koss & Nielsen, 1997), suggest that the technology may now be ready to be used in this manner.

The purpose of this chapter is to assess the ability of the state-of-the-art in modeling technology to solve pressing problems in the engineering of complex systems. On the one hand, we seek to identify the 'low hanging fruit' – obvious and immediate applications of human modeling technology. On the other hand, we seek to identify areas where additional research and development is needed, and from those to identify areas that have a high payoff potential.

The remainder of the introduction briefly establishes an historical context of human performance modeling research. The second section creates a taxonomy of modeling technologies, which will be used throughout the rest of the chapter. The third section focuses on the application of human modeling techniques to support the design and evaluation of a complex system. Specifically, the section addresses both how human models could be incorporated to support the relatively common practice of simulation-based design evaluation, and how human modeling technologies might be applied to support the generation of design concepts. The validation and maturity of the human performance models themselves are also discussed in this section. However, these are large topics in there own right, and a full discussion is beyond the scope of this chapter. The fourth section focuses on the application of human modeling techniques to support the operation of complex systems. These applications typically involve embedding the model(s) in the operational system to provide support for activities such as decision-making, embedded training, and task or workload management. The final section summarizes the current capabilities and highlights and identifies the needs and opportunities for longer-term research, based on potential application payoff.

Historical Development and Application of Human Modeling

While work on computer simulation techniques did not begin until the emergence of commercial computers in the 1950s, many of the basic quantitative models of human performance used in those simulations were actually developed much earlier, often in university research laboratories. Modern techniques for the analysis and modeling of cognitive performance can be traced back at least to the seminal work of Donders in the Netherlands in the mid-19th century. Current models for basic processes in visual and auditory perception can be related back to the late 19th century research of Helmholz, Hering, Weber, and Fechner. Work from the early part of this century on the dynamics of eye movements (Dodge & Cline, 1901) and hand movements (Brown & Slater-Hammel, 1949) continues to be relevant in contemporary models. Indeed, a very large literature has developed on models for many distinct aspects of human performance, ranging from basic features of sensation, through many kinds of perception, motor activity, information processing, and decision making. Fortunately, recent efforts have also addressed the compilation and indexing of information about this rapidly growing body of human performance models (e.g. Boff, Kaufman & Thomas, 1986).

Individual Level Human Performance Model

Formulation of the first computer simulations of human performance began in the late 1950s along two very different tracks. Beginning in the late 1950s was the work of Siegel and Wolf (1962, 1969) to construct computer simulations of the performance of individuals and teams in operating complex military systems. The Siegel-Wolf models represent the first of the class of task network models that describe the dynamics of task performance in terms of the sequence and timing of subtasks. This approach was later translated into a general purpose modeling tool for the Air Force in the form of a modeling system designated as SAINT (Systems Analysis of Integrated Networks of Tasks), which was subsequently adapted for micro-computer use and commercially distributed under the name MicroSAINT (Laughery, 1998). These task network models have been applied in the analysis of a large number and variety of contexts and have been shown to provide a very efficient means for simulating such large, complex systems.

Around the same time, Allen Newell and his colleagues at Carnegie-Mellon University were taking a different approach to understanding and modeling human performance, beginning with the concepts of the logic theory machine and the general problem solver and leading through the mainstream development of artificial intelligence and cognitive modeling to the current GOMS (Card, Moran & Newell, 1983) and SOAR techniques. These models and techniques focus heavily on describing how performance is derived from knowledge, which in turn must be represented in fine detail in order to support performance predictions.

Initially, a modeler had to choose between these two fundamentally different approaches, with task network models requiring the user to provide estimates for the times and accuracies of behaviors at the finest level of model representation, and knowledge-based models affording the promise of obtaining basic performance data from general perceptual, motor, and cognitive models (Laughery & Corker, 1997). However, it wasn’t long before modelers began to look for ways to integrate the two approaches. The first general purpose modeling tool which sought to integrate the task network and knowledge-based features was the Navy's Human Operator Simulator (HOS) which was initially developed in the early 1970s (Lane, Strieb, Glenn & Wherry, 1981). This tool provided a mechanism for representing the detailed configuration of displays and controls in a crewstation, the task procedures by which the user operates the crewstation, and the component "micro-models" by which the human interacts with the displays, controls, and related information.

SOAR was unveiled by Newell and colleagues in the mid-1980s (Laird, Newell & Rosenbloom, 1987) as the first major offering of a general architecture for human cognitive performance, promising to provide the capability to represent the complete range of human behavior from a knowledge-based perspective. In addition to representing both declarative and procedural knowledge, SOAR also incorporates a learning mechanism based on the use of various general problem-solving strategies and the "chunking" of successful solutions. SOAR has been widely used in university laboratories around the world and has recently been demonstrated in a large-scale application for computer-generated forces (Laird, Coulter, Jones, Kenny, Koss & Nielsen, 1997).

Since the emergence of SOAR, several other candidates have been developed as general computational architectures for representation of human cognition, such as ACT-R, EPIC, and COGNET. ACT-R was developed by Anderson (1993) as a model for higher level cognition, such as in problem solving tasks, with a principal focus on the investigation of mechanisms of learning. EPIC was developed by Kieras and Meyer (1997) to provide a detailed representation of how human task performance is dictated by the constraints imposed by perceptual, motor, and cognitive abilities, with special emphasis on defining how perceptual and motor activities interweave with other aspects of cognition. COGNET was developed by Zachary and colleagues (Zachary, Ryder, Ross & Weiland, 1992; Zachary, Le Mentec & Ryder, 1996) as a model of expert-level problem solving and task performance in real time, multiple-task environments. Pew and Mavor (1998) reviewed ACT-R, COGNET, EPIC, MicroSaint, SOAR, and several other integrative modeling architectures in detail, in the context of military simulations of human behavior.

Group and Team Level Human Performance Models

In addition to these general models and architectures for individual human cognition and performance, other types of models represent and study the performance of multiple individuals operating as teams or larger organizations.

Unlike the psychological models, which have focused on individual cognitive and behavioral processes, these social science models have tended to focus on the structure of human interactions and social systems. Whereas individual-level models have focused on processes, these group and team models have been primarily concerned with the role of constraints and structure on processes, rather than on underlying mechanisms. However, similar to the psychologically based individual-level models, the team and organizational models of current interest also come (primarily) from two distinct lines of research, which can be termed the micro and macro approaches.

The micro approaches focus on modeling interaction upward from the (atomic) level of the individual dyadic relationships. This work has been heavily mathematical and influenced by the early work of Harrary, White and others who employed the mathematics of graph theory as a framework for modeling the networks of relationships among people. During the 1960s and 1970s, very sophisticated models were derived from this “social network as graph” simile, (e.g. see White, Boorman and Breiger, 1976, or Zachary, 1977, for very different network modeling approaches), among others. This research showed how the collection of individual dyads that makes up teams, groups, and organizations has a deep structure that could often be seen to have clear effects (or at least reflections) in the activities and processes that occurred within these social units. At the same time, it became increasingly clear that collection of the data on individual dyads was a costly and time-consuming process, and a series of experimental studies, termed the “informant accuracy studies”, demonstrated the difficulties in measuring the dyadic structure of social networks on an indirect (i.e. non-observational) basis.[2] These measurement and data collection problems have continued to hamper broader development and application of the field to the present day.

The macro approach focuses on broader processes in social groups, particularly large social units such as organizations, cities, societies, etc., without reference to individuals and/or their unitary dyadic relationships. This tradition achieved prominence with the work of Jay Forrester at MIT (Forrester, 1971) and the Club of Rome (Meadows, Meadows, Randers & Behrens, 1972), who began to model the behavior of complex systems as the dynamic interactions of multiple complex sets of constraints and underlying relationships. Formulated as systems of difference and/or differential equations, the resulting system dynamics were instrumental in demonstrating how underlying relationships had long-lasting and subtle effects on the long-term behavior of the larger systems. Just as the micromodels of group structure showed how the pattern or relationships in a network led to higher level structures and constrained group processes, the dynamic models of the macromodelers demonstrated how a constant set of underlying relationships could give rise to a broad range of complex and varying processes through time.

A third, and now dominant, approach has arisen from a combination of the micro and macro approaches (see Pew & Mavor, 1998: Chapter 10 for a more detailed review). In this approach, which can be termed the agent-based approach, the individuals within a team or organization are represented in simplified fashion as agents, interconnected by networks of command, control, and communication relationships. Typically, both the agents and the network in which they are embedded are dynamic, leading both to learning and organizational changes over time.[3] A critical issue in this approach is the degree of cognitive sophistication given to the agents, and the degree of organizational sophistication of the network in which they are embedded.

Use of Human Models in System Engineering

Just as the research into modeling human behavior is not new, the current attempts to apply human models to complex systems engineering are not the first of their kind either. In the early 1970s, for example, the Navy initiated the CAFES (Computer Aided Function Evaluation Systems) program to develop simulation-based tools for assessment of workload, function allocation, and anthropometric accommodation in aircraft cockpits (Hutchins, 1974). This was followed by the Air Force CADET (Computer Aided Design and Evaluation Tools) program (Connelly, 1984) and then another larger scale Air Force program on Cockpit Automation Technology (CAT) (McNeese, Warren & Woodson, 1985). In the mid 1980s, the Army began the development of a collection of simulation-based tools to support the methods of the MANPRINT program, since designated as the HARDMAN-III tools (see Risser & Berger, 1984). At about the same time, the Navy started the Advanced Technology Crew Station (ATCS) program to develop and demonstrate the use of simulation and computer-based tools in design and development of new aircraft cockpits. More recent programs in this vein include the Air Force's OASIS program and NASA's MIDAS program. And there have been several systematic reviews and conferences in this area over the past few decades (e.g. Moraal & Kraiss, 1981; McMillan, Beevis, Salas, Strub, Sutton & Van Breda, 1989; Elkind, Card, Hochberg & Huey, 1990; Baron, Kruser & Huey, 1990). Of course, the technologies of human performance modeling and simulation are not stationary targets, and their rapid development over recent years warrant periodic reexamination of status and development needs.

In general, these DoD programs did not develop the basic human performance models, but rather sought to adapt existing models and techniques in order to produce practical tools to aid system designers and developers. For example, several of the Army's HARDMAN-III tools used the same basic human performance model (variants of MicroSAINT) for distinct applications to determine system performance requirements, to evaluate workload for a team of system operators, and to assess the effects of individual differences. While these and other programs were successful at applying human performance models to some specific 'target' system or concept, they have, so far, failed at the more global task of making human modeling an accepted and standard component of complex systems engineering within DoD.

THE TECHNOLOGY BASE

Human cognition and behavior are modeled through a variety of approaches and technologies. Some of these focus on individual components of human performance (e.g. models of visual target detection), while others focus on the integration of components at the architectural level (e.g. executable cognitive architectures). The various ways in which human models can be used in complex systems design, evaluation and operation do not inherently favor either component or integrative approaches. For example, in designing a sensor display to optimize human target detection, a model of visual detection may be completely sufficient. A more integrative model that included anthropometry, auditory processing, cognitive planning, and problem solving, etc., might well be 'overkill' and too cumbersome to justify its use. On the other hand, applications such as large-scale system simulation or distributed training exercises may require highly integrated modeling approaches. To facilitate a more systematic mapping of the human modeling technology base onto the various opportunities and requirements in the complex systems domain, a simple taxonomy of the existing technology base is presented below.

The taxonomy uses two broad groupings – models and modeling techniques. *Models* includes complete formulations (or families of them) that attempt to describe, predict, or prescribe aspects of human competence or performance, either component-wise or integrative. *Modeling techniques* includes computation, mathematical, or methodological formulations that have been used to build models of human competence/performance or to apply human models to system design, operation, or evaluation problems. They differ from the first grouping in that modeling techniques are more general purpose tools for modeling (in mathematics, physics, computer science, etc.) that have been sometimes used to represent human cognition or behavior. These modeling techniques do not embody specific psychological or sociological theories, nor were they developed specifically for the purpose of human modeling.

Following the discussion of the various categories of models and modeling techniques, the next sections will consider the use of this human modeling technology in the complex system design/evaluation process, and in the complex system operation process.

Classes of Human Models

Human models, as defined, may be classified into 18 categories, which span the range from highly specialized component models (e.g. perceptual models) to highly integrative representations (e.g. computational cognitive models). These categories are (in alphabetical order):

(1) *Closed form component models*. Self-contained mathematical formulations that represent some component aspect of human performance as a self-contained closed form mathematical relationship rather than any common application or set of underlying terms. Fitts' Law (Fitts & Peterson, 1964) is a classical example of this type of model.

(2) *Computational cognitive models*. Integrative models of human cognition, perception, sensation, motor action and knowledge that embody a principled underlying theory or framework for human information processing. This class includes items such as ACT-R (Anderson, 1993), COGNET (Zachary, Le Mentec & Ryder 1996), EPIC (Kieras & Meyer, 1997), and SOAR (Laird, Newell, and Rosenbloom, 1987), among others. These models capture human knowledge in a symbolic form and allow behavior and cognition to be generated as a result of a symbolic computation process.

(3) *Critical decision models*. Domain-specific models that capture and represent the logic and situational relationships that underlie decision-making in that specific domain (Klein, Calderwood & MacGregor, 1989). These descriptive models typically focus on extracting knowledge from expert decision-makers and representing it in a combined prose/graphical notation.

(4) *Decision theory models*. A broad family of normative models that represent human behavior in choice-among-alternatives situations. Generally drawing on the terms of game theory (Von Neuman & Morgenstern, 1947), the choice process is represented in terms of outcomes, outcome-utilities, and underlying distributions of input states, together with varying parameters that include factors such as subjective value, risk, and risk preference, etc.

(5) *Finite state models*. A computational framework for relating inputs to behaviors using the notions of internal information states and the mathematics of finite state automata.

(6) *GOMS (Goals, Operators, Methods, Selection Rules)*. Domain-specific models of the knowledge used in human-computer interaction that are developed using a notation created originally by Card, Moran and Newell (1983). GOMS models decompose primarily procedural knowledge into goal hierarchies, which are conditionally related to states of the interactive process and interface.

(7) *Group Training Models*. Team-level representations that could either relate collective (i.e. team) training to performance or to models that relate performance to training requirements.

(8) *Human reliability analysis*. Various models for estimating the likelihood of errors occurring on complex tasks as a function of elemental task error probabilities and other factors. These techniques are used, for example, to estimate the impact of human errors on system performance and to evaluate system designs and recommend improvements (Czaja, 1997).

(9) *Link models of anthropometry and movement*. Numerical/graphical models that represent the human body and its capability for movement and vision through an articulated set of 'links' which themselves represent lower-level body components (e.g. fingers, arms, torso, etc.). In the most sophisticated of these models, the links are represented not as simple lines but rather as complexly-interrelated solids.

(10) *Network models*. Models in which individuals, teams, groups, and/or organizations are represented as nodes and their dyadic relationships as links in a graph-theoretic mathematical representation. The degree of complexity in the dyadic representation varies greatly, and the mathematics often make recourse to higher-level systems computed on the raw graphs.

(11) *Optimal control models*. Mathematical/computational representation of decision making and/or adaptive behavior in an uncertain environment, based on the underlying mathematics of optimal control theory (Kalman, Falb & Arbib, 1969). These models include an internal model of the external world which is used to select behavior in the world based on current input, current output, current internal state, and various filters.

(12) *Perceptual models*. Various descriptions of sensation and perception for any of the human sense modalities, though primarily for vision and hearing. The possibilities range from models for detection of simple stimuli, through interpretation of complex patterns.

(13) *Recognition Primed Decision Making (RPD)*. Models of domain-specific expert human decision making that are developed using an underlying RPD framework. This framework asserts that decisions emerge from an (implicit or explicit) situation assessment process that maps situational features and understanding onto appropriate decision-options. Developed by Klein (1989), the RPD framework is related to more general modeling techniques of case-based reasoning.

(14) *Signal Detection Theory.* Models for determining the probability of detection of a stimulus (which may be very simple or quite complex) based on the assumption that the signal (i.e. stimulus) and the background noise can each be characterized by normal distributions along relevant psychological dimensions (Green & Swets, 1966).

(15) *Task Hierarchies*. Domain-specific models of task performance achieved through a conventional task analysis that simply decomposes a job into a hierarchy of tasks and subtasks.

(16) *Task networks*. Domain-specific or job-specific models which decompose human activity in the domain or job into a series of interconnected tasks, where tasks are represented as transformations of some state/variable vector, and task-connections represented as potential task transitions, often conditioned by probabilistic factors. Originating from the SAINT methodology (Wortman, Duket, Seifert, Hann & Chubb, 1978) the network is often used to simulate the behavior of the system as well.

(17) *Training taxonomy.* A structural model of generic tasks in terms of the underlying skill requirements so as to support analyses of training effects on skill retention and, hence, task performance (see Swezey & Llaneras, 1997).

(18) *Workload Models*. Models which represent residual human work or information processing capability in specific context. Typically workload models are based on an underling theory, such as Wicken's (1980) Multiple Resource Theory which represents information processing and capacity as the result of interactions among multiple lower-level capabilities or resources.

These eighteen categories represent only broad classes of human models, and even then only a subset of human models that appear particularly relevant to the complex system design and operation problems. The functional definition of the classes allows many vastly different models to be included within a given class.

Regardless of the class to which a model belongs, however, it can be described along four key dimensions:

(1) *Model goals*. Models that attempt to describe the regularities of human behavior or cognition within a specific setting that are captured in empirical data are called *descriptive models*. Other models, called *predictive models*, attempt to predict human behavior/cognition in a range of (future or hypothetical) situations described in terms of empirical data. Finally, there are *prescriptive models* which attempt to prescribe what a person should do/think in a given (future or hypothetical) situation.
(2) *Architectural focus*. A given model may focus on individual architectural aspects of human capabilities and behavior, including: (a) the characteristics of the *physical* body and the ability of the body to move and act in a given physical environment, (b) the processes of registering information from the external environment (*sensation*) and/or internalizing registered information into a form/representation used for internal information processing (*perception*), (c) the representation of information in the mind and the process of manipulating that information to yield complex behaviors such as reasoning, decision-making, or planning (*information processing*), or (d) the process by which the person acquires information and knowledge about the external world and external processes, and internalizes this information and knowledge on a long-term basis for use in future information processing (*learning*). There are also *integrative* models, which attempt to integrate all of multiple components into a single model of human cognition and behavior.
(3) *Predictive focus*. Those models that seek to predict human behavior can differ on the aspects of human capability that they seek to predict. The most basic kind of predictive focus is on the *outcome*, or specific behaviors that a person will take and the outcomes of those behaviors in a given scenario or situation. There are also *time*-based models, which seek to predict the time it takes for activities and/or actions to be performed, and *accuracy*-based models, which seek to predict the accuracy with which activities and/or actions will be performed. *Workload* models seek to predict the workload, either total or by-component, associated with performing specific actions and/or activities. Some models seek to predict the *situation awareness* that a person will have of the internal and external situation at a specific point in a specific scenario. Finally, some models may attempt to predict the internal *structure* or organization of an activity or process.
(4) *Level of modeling*. Finally, models can address different levels of human activity, either the *individual* level, or the *organizational* level, which is taken here to include models of dyads, teams, organizations, etc. (any unit above the individual level).

There are other possible dimensions along which models could be compared, but these four are discussed here because they are relevant for the issues and assessments made in later application sections of this chapter. Table 1 shows the eighteen model categories evaluated on their ability to support the construction of models with different values along these four dimensions.

Technology Maturity

The range of human modeling technology discussed above is broad. These various technologies are not at the same stage of development, nor are they at the same point of readiness for practical application to complex systems engineering. Model maturity can be rated according to three categories:

- Applicable Now – the modeling type/technique has been used successfully or is judged to offer sufficient potential for successful application to some aspect of system design today.
- Applicable in the Short term – the modeling type/technique is at a state where several years of focused research and development coupled with supporting model validation research would probably make it applicable to system design.
- Applicable in the Long term – the technique, while offering promise, is not yet at a state of maturity where a usable, practical technique can be envisioned, but the technique still merits further research and development.

In general, all the model classes listed in Table 1 are applicable now, with two exceptions. Both Group Training Models and Training Taxonomies are in early stages of development and are applicable only in the long term. Also, Computational Cognitive Models are rapidly evolving, and substantial new capabilities are likely to be available in the short term.

Validation of Human Models

In various engineering communities, there are standards and processes for model verification, validation, and accreditation. Of course, the behavioral sciences community has independently established its own expectations as to what is meant by the term validity. These communities have not been well coordinated, and a plan for assessing human performance model validity that addresses concerns of both communities is an immediate concern.

Before any model can be used for engineering decision making, the model's validity and utility must be established. Three general criteria are commonly used to assess the validity of a given human model. *Predictive/concurrent*

Table 1. Model Classes Rated on Key Comparative Dimensions.

Model Class	Goals	Architectural Focus	Predictive Focus	Level
Closed form component models (e.g. Fitts' Law)	descriptive, predictive	various	outcome, time, accuracy	both
Computational cognitive models	descriptive, prescriptive, predictive	integrative	outcome, time, accuracy, situation awareness, workload	individual
Critical decision models	descriptive	information processing	outcome, situation awareness	individual
Decision theory models	descriptive, prescriptive	information processing	outcome	individual
Finite state models	descriptive, predictive	integrative	outcome, time, situation awareness	individual
GOMS	descriptive, prescriptive, predictive	information processing	outcome	individual
Group training models	descriptive	learning/training	outcome	organizational
Human reliability analysis (HRA)	descriptive, predictive	physical, perception	time, accuracy	individual
Link models of anthropometry & movement	descriptive, predictive	physical	outcome	individual
Network models	descriptive	integrative	outcome, structure	organizational
Optimal control models	descriptive, predictive	integrative	outcome, time, accuracy, situation awareness	individual
Perceptual models	descriptive, predictive	perception	outcome	individual

Table 1. Continued.

Recognition Primed Decision-making (RPD)	descriptive, predictive	information processing	outcome, situation awareness	individual
Signal Detection Theory	descriptive, predictive	perception.	outcome	individual
Task hierarchies	descriptive, prescriptive	behavior	process	individual
Task network	descriptive, predictive	behavior	outcome, time, accuracy, situation awareness, workload	both
Training taxonomy	prescriptive	learning/training	structure	individual
Workload	descriptive, predictive	information processing	workload	individual

validation is established by comparing the predictions of the model to actual performance data. *Construct validation* is established by demonstrating that the underlying constructs and model components are valid. *Face validation* is established by submitting the model for review by experts to assess apparent validity. These three types of validation are presented in the order of the power of the statement of validation (predictive validation makes a much stronger argument than face validation). This typically also translates into cost (e.g. predictive validation studies involving sufficient numbers and types of subjects will be far more expensive than the less-strong face validation studies).

Table 2 presents a summary of the types of validation studies that have been conducted for the human performance modeling types introduced earlier in this section. There is one important caveat regarding Table 2. The entries in the table reflect not that the modeling approach (e.g. task network modeling) has been validated, but rather that models built of specific systems using this approach have been validated. Overall approaches can not be validated per se except to the extent that the underlying scientific theory they embody or formalize can be validated. Validating a particular model (application)

Table 2. Estimated Level of Validation Studies on Model Types.

Modeling Class	Validation Studies Conducted
Closed form component models (e.g. Fitts' Law)*	Face/construct/predictive
Computational cognitive models*	Face/construct/predictive
Critical decision models	Face
Decision theory models†	Face/construct/predictive
Finite state models	Face/construct/predictive
GOMS*	Face/construct/predictive
Group training models	Face
Human Reliability Analysis	Face/construct
Link models of anthropometry & movement	Face/construct/predictive
Network models*	Face/construct/predictive
Optimal Control Models	Face/construct/predictive
Perceptual Models	Face/construct/predictive
Recognition Primed Decision-making (RPD)*	Face/construct/predictive
Signal Detection Theory	Face/construct/predictive
Task hierarchies*	Face
Task network*	Face/construct/predictive
Training Taxonomy	Face
Workload*	Face/construct/predictive

* Varies with individual model in this class. † Models in this class have been shown to have very constrained prediction validity

example, on the other hand, is more constrained and more possible. At the same time, if a model type or approach has been demonstrated to be valid in specific applications, it is easier to make the argument for its validity in other applications.

While the goal of validating human performance models by predictive validation (the highest standard, as noted above) is commendable, this is a high and often unachievable goal. Human performance is inherently variable which means to get stable measures of performance against which to compare model predictions requires large numbers of subjects. Also, human performance research in complex systems is expensive and, particularly for evolving systems, often impossible. It is therefore the case that lower standards than predictive validity may be adequate. In fact, typical military verification, validation, and accreditation (VV&A) does not consistently require predictive validity as part of the validation process. In addition, it is probably the case that different model types may have different requirements. For example, models such as anthropometric or perceptual models may require predictive validity to be established (in either a general or situation-specific context) before they can be reasonably used in system validation, while for other model types (particularly those that are highly descriptive, such as Critical Decision Analysis, GOMS, task network representation) the concept of predictive validation may be less meaningful and not required.

In a broader context, however, the assessment of model validity is in fact subject to different standards and processes. Specifically, academic and scientific research utilizes one set of principles and standards for establishing the validity of models, particularly as embodiments of scientific theory, while engineers and simulation analysts, particularly in the military, have very different processes for VV&A of models for use in military analysis. In the narrow sense of specific models being used in engineering applications, it is reasonable that established requirements of the military VV&A process should form a sufficient test of validity in most cases (subject to the arguments discussed in the preceding paragraph). This suggests that the human performance modeling and systems engineering communities rely on construct and face validation where these may be sufficient conditions. This will allow experience to be gained in using the models to build actual systems and will ultimately lead to empirical usage data that can be used to compare model-based predictions to actual outcomes, and thus lay the foundation for broader predictive validation.

Finally, *model usefulness* should be considered as relevant to model validity. Four dimensions related to model usefulness can be identified. First, essential to any determination of a model's utility is defining when the model is appropriate and, by inference, when it is not. Factors to be considered include the underlying assumptions and the situations under which the assumptions are met (for example, a visual model may assume high levels of ambient light, making it inapplicable in dimly lit or night-vision situations), the type of human performance outputs predicted by a model (e.g. models that do not predict performance time would be inappropriate for evaluating data through-put rates), and known limitations from prior validation studies. Models with well-defined boundaries will be more useful because their applicability will be easier to assess. Second, for a human performance modeling tool to be useful, it must be usable. During the assessment of model usefulness for a purpose, one must consider what is required of the target user in order for the tool to be usable. A modeling tool that takes months to use when only weeks are available is not useful when there are severe schedule constraints. A modeling tool that requires several years of education or can only be used by highly specialized people is limited to applications where there is a high potential payoff.

In addition to being usable, a model (or just the use of modeling) may add value beyond the narrow confines of the modeling application. For example, the analysis required to build the model may lead to other engineering, design, or conceptual insights, or results that have nothing to do with the actual model application. This kind of "value added" contributes to a model's overall usefulness. Finally, an increasingly important component of model usefulness is its interoperability, or the extent to which the model can be linked to other models of hardware, software, and even other types of human performance models. The easier it will be to get a human model into an integrated modeling environment, the more useful the model will be.

Relevant Modeling Techniques

The state of the art of human modeling has certainly not progressed to the point that there is a well-defined collection of models that can be used in all commonly occurring situations. There are still many cases where a human model has to be cobbled together ad hoc from lower-level building blocks for a specific, unique situation. A number of these "building blocks", or modeling techniques, are presented below and discussed in terms of five high-level categories, as shown in Table 3.

Knowledge Representation Techniques are methods that have been developed, typically in artificial intelligence and cognitive science, to capture and represent human knowledge in symbolic form for use by computational models of reasoning and problem solving. Three commonly used techniques for knowledge representation include blackboards (Carver & Lesser, 1994),

Table 3. Modeling Techniques Applicable to Human Model Development.

Modeling Function	Modeling Technique
Knowledge Representation	Blackboards
	Production Rules
	Semantic Nets
Decision Making and Reasoning	Bayesian Inference
	Fuzzy Logic
	Hybrid Logic
	Neural Nets
	Hidden Markov Models
	Case-based Reasoning
Design/Analysis Techniques	Early comparability analysis
	Simulation-based tutoring and authoring
	Simulation based walk-through
	Statistical models (descriptive)
	Statistical models (interpretive)
Natural Language Processing Techniques	(various methods)
General Purpose Simulation	Monte-Carlo methods
	Dynamical models

production rules (Anderson, 1976) and semantic nets (Collins & Quillian, 1969). Blackboards are hierarchical representations of primarily declarative knowledge that are used in opportunistic reasoning and open problem solving systems. Production rules, on the other hand, are abstracted representations of atomic if/then propositions, which can be applied inductively and deductively to specific sets of facts to yield new facts and inferences. This is the form of representation typically used in expert systems. Finally, semantic nets are representations of facts and the semantic relationships among them (e.g. 'part of', 'kind of', and other more domain-specific forms), typically developed using graph representations.

Reasoning Techniques are methods that have been developed to allow the computational representation of problem-solving, planning, decision-making or other reasoning processes. These methods were not developed as models of human performance per se, but have sometimes been used to develop human cognitive/behavioral models. Commonly used techniques include Bayesian inference, fuzzy logic, neural networks, hidden Markov models and case based reasoning. Bayesian inference techniques are mathematical and computational methods to permit reasoning about uncertainty based on the underpinning of Bayes rule (that the probabilities of all disjoint events sum to unity). A particular manipulation of Bayes formula for conditional probability is of particular importance, as it allows reasoning about sequences of discrete events (i.e. data) in which (prior) estimates of underlying distributions are updated following each observed event (to yield a posterior distribution). This process has been used to model diagnostic inference processes. Fuzzy logic techniques, on the other hand, are mathematical and computational methods to permit reasoning about uncertainty without requiring the mathematical constraint of Bayes rule. Originally developed by Zadeh (1965) as an artificial intelligence method, it has been offered as an alternative (i.e. non-Bayesian) way of modeling human reasoning and decision-making under uncertainty. (Note: there are methods that incorporate multiple aspects of logic and both Bayesian and non-Bayesian inference methods into a single framework, and these are typically called hybrid logic approaches.) Neural nets are methods that simulate the processing of information by a distributed and highly interconnected network of (typically simulated) simple information processing devices analogous to individual neurons (see Rumelhart, McClelland et al., 1986). Hidden Markov models are processing algorithms, often used in language processing models, that provide discrete stochastic state-based representations of sequential relationships (i.e. Markov models) organized in a hierarchical manner, so that what appear to be states at one level are actually complete Markov processing models at a lower level of detail. Finally, case based reasoning methods categorize problem solving processes into alternative strategies that are indexed by archetypal examples called ‘cases’, and allow individual problem-solving examples to be mapped into appropriate problem-solving processes according to their similarity to the various case archetypes.

Design/Analysis Techniques are methods that have been developed to support the process of designing systems or components of systems. In many cases, though not all, these techniques either can interface with models of humans or can produce situation-specific human models. For example, one standard engineering technique, early comparability analysis, is a method in which design situations or design candidates are compared with other existing implemented systems or subsystems as a basis for estimating characteristics such as manning requirements, training requirements, complexity, etc. Another common technique, simulation can be used in a number of ways, including simulation based tutoring and authoring, a method by which rapidly constructed simulations are used to capture procedural knowledge and support the interactive tutoring of novices in those procedures. Simulation based walk-through is a method in which a simulation of a system or system design is used as a basis for visualizing or ‘walking through’ that design, which is typically used in conjunction with virtual prototyping technology that provides the highfidelity simulation of the system. Finally, there are a series of classical, statistical methods which either describe the characteristics of specific populations in terms of distributions of random variables (descriptive), or model the relationships among (interpretive) groups of independent or interdependent random variables (e.g. among arm, leg, trunk length in contributing to individual height in a human population).

Natural Language Processing Techniques are a family of primarily computational techniques that model the human ability to process language in written (i.e. text) form. These techniques can allow other types of human models to interact with real humans in simulations or games, or can be used to provide model users with natural language access to the functionality of other models or techniques.

General purpose simulation methods are computer science and operations research techniques that are used to simulate systems of all types, not just human systems. Although there are many ways to categorize these methods, Zachary (1986) uses two orthogonal dimensions (mechanical vs. analytical and stochastic vs. deterministic) to create four categories. Mechanical methods are those that decompose a process into discrete events and processes and mechanistically simulate the system's behavior in terms of those events and processes. Often called discrete event simulation, mechanical models can be either stochastic or deterministic in the way they process through the event/process space. The class of human performance models termed "task networks" represents the application of mechanistic simulation methods to human modeling. Analytical methods are those that describe a system in terms of the underlying relationships, typically expressed in mathematical terms, that describe either input/output transformations or time-dependent relationships. Dynamical models, on the other hand, are analytical models which represent the changes in a system through time as a set of simultaneous differential (continuous change) or difference (discrete change) equations which are functions of time, or by dynamically interacting agents. These models can also be either stochastic or deterministic, although deterministic representations are more tractable and more common.

USING HUMAN MODELS IN SYSTEM DESIGN

Today, it is standard practice to use models of system hardware and software during concept exploration, preliminary design, and full-scale engineering development of a complex system. One of the functions of these models is to ensure that the system performance objectives, whatever they may be, are satisfied by the design. Additionally, the models ensure that the performance objectives are met within any defined design constraints, such as weight, power-consumption, etc. This use of models reflects a (often implicit) focus on the system *sans* humans, as if the human role was unimportant, imponderable, or both. Increasingly, though, this view is being called into question, often as the result of designs that have proven unusable, untrainable, or unstaffable. A new view is emerging in which the performance of the *integrated* system – hardware plus software plus people – is the explicit focus of performance objectives. In other words, the human role is explicitly viewed as important and even critical. And as the discussion above shows, the modeling of this human role is certainly no longer imponderable; a wide range of models and modeling tools exist. The question that arises, then, is what types of models should be used for what purposes in the design process?

We begin with a brief overview of the design process for complex systems, and from that discuss two major uses of human models – one traditional, and one non-traditional. The traditional application is the use of human models to validate design concepts, by assessing whether explicit design concepts can meet their performance requirements given the capabilities and limitations of their human components. This is a traditional application because the human model is used in exactly the same way that hardware and software models are now used, and there is a growing history of success at using human performance models in this way. The non-traditional application is the use of human models to support the synthesis and concept generation activities associated with the design process. This is a relatively new and somewhat exploratory area of application for human performance models.

The Complex System Design Process

The design of complex systems is a main concern of the discipline of system engineering. As any engineering discipline, it is the subject of ongoing discussion, research, and debate. There are several standards produced by various organizations, e.g. Institute of Electrical and Electronic Engineers (IEEE Standard 1220–1998), International Committee of System Engineering (INCOSE Systems Engineering Handbook), Government Electronic and Information Technology Association (EIA standard 632), and the U.S. Department of Defense (Directives 5000.1 and 5000.2-R), and competing theories of how the process should be done (cf., Blanchard & Fabrycky, 1998). Thus, there is no single commonly agreed standard for how a complex system should be designed. Still, it is possible to identify the main features from most commonly applied approaches, and distill them into an abstract representation that can guide the consideration of where and how human models might fit in.

Most approaches fall along a continuum from what is called ‘waterfall’ to what is called ‘spiral’. Both processes proceed through a well-defined series of stages, generating concepts, evaluating them, selecting the best, developing detailed designs from it, and implementing and testing those designs. Waterfall approaches tend to go through this process only one time, building systematically from coarse granularity to fine granularity in the design and implementation. Spiral approaches tend to proceed through the process many times, each time creating a more-complete design or implementation, correcting problems from past cycles, and adding new detail in areas previously ignored. The core steps of both approaches can be seen in Fig. 1.

As a general overview, the system design process in Fig. 1 has three fundamental stages. Initially, designers deal with mission analysis and requirements analysis (WHY is this type of system needed?). Next, designers deal with functional analysis and logical design (WHAT will the system have to do?). Finally, designers deal with implementation details associated with a particular design, such as allocating functions and resources (HOW will the system accomplish that?). At each stage, there are two types of activities. *Design synthesis* activities are the tasks a designer engages in to come up with a potential requirement/logical design/detailed design. Design synthesis typically involves a high degree of art and creativity. *Design analysis* activities include systematic, thorough, and rigorous attempts to evaluate a potential requirement/logical design/detailed design. The overall cycle can be iterative in nature, as in a spiral process, or recursive as in a waterfall process, where

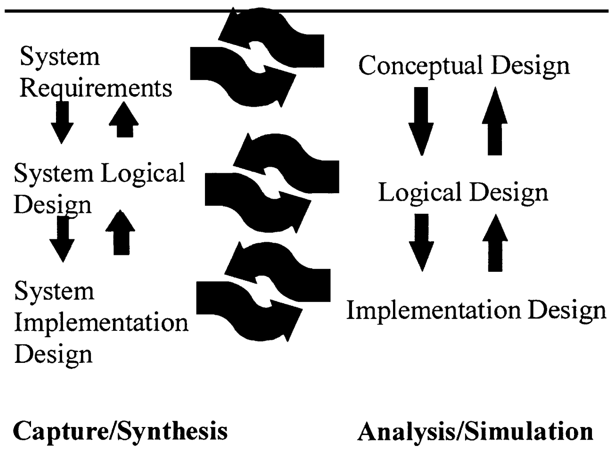

Fig. 1. Abstracted System Design Process.

systems are broken into subsystems, which are broken into sub-subsystems, etc. Information needs to roll up and down the hierarchy, explicitly identifying the implications of a design implementation of a subsystem on a variety of measures, including cost, impact on other subsystems, ability to meet the mission needs, etc.

Of course, complex systems are rarely, if ever, built entirely de novo. Certain components are typically adapted or incorporated from pre-existing systems on a legacy basis, to limit the complexity of the design process and to limit the developmental cost of the final system. The amount of novelty in the system can and does vary, however, and this strongly affects the design process. In general, a given design problem can also be seen as falling on a continuum from evolutionary (involving an incremental moderate modification to an existing system) to revolutionary (involving a total departure from the design of existing systems). Few systems are totally revolutionary, as the cost of making everything new is prohibitive, yet few are also totally evolutionary, as new designs are rarely warranted without substantive change.

More revolutionary design problems, though, can lead to more open-ended design processes. An interesting case in point is the Navy's goal of a 70% manning reduction on future Navy ships (Carnevale, Bost, Hamburger, Bush & Malone, 1998). The realities of cost and financing dictate that it is not pragmatically possible to start designing this ship with a totally clean sheet of paper. The extreme requirement for manpower reduction, though, may require a design that is "revolutionary" in that it includes all aspects of design – function automation, consolidation, elimination, and simplification – and considers a wide variety of factors including, but not limited to, human factors. For example, improving the human-computer interface (HCI) of a system could reduce manpower requirements by allowing fewer people to accomplish more work. On the other hand, a better paint needs less repainting, which also reduces manpower requirements. It is interesting to note the revolutionary requirements may not be unidirectional. Achieving global manning reductions on this future ship may not require proportional reductions in each area, and may even require increasing manning in some areas. For example, as a consequence of the increased automation during the 'Smart Ship' project, a training department was added to the ship (see Giffen, 1997).

Revolutionary system requirements and design processes can also conflict with the more traditional life-cycle model of complex systems, in which an initial version is implemented and fielded, and then incrementally modified over a long period of time. This process minimizes the marginal cost of any modification, as (in theory) only the new functionality must be designed, implemented and integrated, rather than a whole new system. However, it has some pernicious effects, particularly with regard to the role of human components. Here again, ship design is a good example. Historically, a ship was viewed as a collection of component systems “riding side-by-side” inside a common hull. When a new weapon or sensor system was added to a ship, it was added by an incremental modification process. The new hardware was designed, squeezed onto the already crowded deck and/or hull, and a minimally invasive ‘stovepipe’ was created to connect this new hardware with the people who would run it and the other systems (power, data, etc.) that would supply it. The result was that a new console, job function, etc. was created for each new incremental component, which frequently required a new human role. Over time, this stovepipe strategy exploded the complexity and manning requirements of a ship, even though at each step it was the most efficient alternative.

This has led to the concept of total system design, in which the interconnections and ramifications of design decisions at all levels are taken into account, not just the marginal effects. The whole system (in this case the ship) is conceptualized as a single complex system, with the roles of humans and software/hardware components designed together. Under this approach, a top-down functional analysis for the entire ship is performed, and then design options with different allocations among hardware, software and human operators are compared and evaluated. Often this involves task-centered design, in which decisions about the design of subsystems, and about human and machine roles and the human-machine interface, are made with reference to the tasks that must be performed by the person and person-machine combination. Total system design and task-centered design both work within the overall framework pictured in Fig. 1 above, but represent customizations needed to appropriately factor in the design of human work and roles.

Human Models to Support Design Analysis

Human models can be used in the system design process in all three of the stages shown on the right-hand side of Fig. 1, primarily as a means of analyzing the design products on the left-hand side of the figure – system requirements, system logical designs, and system implementation designs. The analysis of requirements involves testing the ‘why’ decisions in the design process, the analysis of logical designs involves testing the ’what’ decisions embedded in the iterative design, and the analysis of implementation designs involves testing of the ‘how’ decisions contained in engineering design.

Historically, the major use of human models has been in the later phases (e.g. Gray, John & Atwood, 1993), but there is a great potential payoff from their use in earlier design phases because they have the greatest impact on the system's likelihood of success and its life-cycle costs. The potential uses of human models in the first two design phases are discussed below together, after which the use in the last design phase is discussed.

Requirements and Logical Design Analysis

These early design phases answer such questions as "What functions will the system be able to perform" and "What general classes of components will be involved in performing these functions"? There are at least four ways in which human models might support these early stages of analysis with current or near-term technology.

The first is identifying stress points and opportunities. Human models can be used to find points in the predecessor system (i.e. the system that the new system will be replacing or supplementing) that are prone to stress or are underutilized. These points provide opportunities for improvement; for example, underutilized human operators could point to requirements for job consolidation opportunities and overloaded human operators could point to requirements for a need for automation. The second use is in identifying organizational requirements. For example, through models, limits on team size and structure can be developed, based upon the need for speed in complex decision making. A third use is in life cycle cost estimation. Human models can be used to develop estimates of the costs of various design requirements and concepts, e.g. the costs of various manpower alternatives. The fourth high-payoff use is in performance budgeting, that is, in using human models to allocate performance requirements to functions and, ultimately, components in a way that achieves overall system performance requirements. For example, this could be used during initial function allocation exploration to assess where the expected functional requirements can be met by each component.

Table 4 presents a summary of the different kinds of human models and modeling techniques that could be used for each of these four types of concept and requirement analyses. In this and similar tables below, the reader is advised the mapping is suggestive only. That is, the authors are aware of applications of this type or believe that applications of that type are possible.

The Use of Human Performance Models to Validate and Evaluate Engineering Designs

In Fig. 1 above, there is a natural flow in the design capture/synthesis process from the 'what' stage, where the establishment of requirements is the main issue, to the 'how' stage, where a determination of how the system will achieve these functions becomes the primary concern. But there is a reverse flow as

Table 4. Models Applicable to Concept and Requirements Analysis.

Model Class	Identify Stress Points and Opportunities	Organizational Design	Life Cycle Cost Estimation	Performance Budgeting
Closed form models			X	
Computational cognitive models	X	X		
Critical decision analysis	X			
Decision theory models	X			
Finite state models	X			
GOMS	X			
Group training models		X	X	
Human reliability analysis (HRA)		X		
Link models of anthropometry & movement	X			
Network models		X		
Optimal control models	X			
Perceptual models	X			
Recognition Primed Decision Making (RPD)	X			
Signal Detection Theory				
Task hierarchies		X		
Task network	X	X		
Training taxonomy			X	
Workload	X			
Modeling Technique				
Knowledge Representation				
Decision-making & Reasoning				
Design/Analysis Techniques	X	X	X	
Natural Language Processing Techniques				
General Purpose Simulation	X	X	X	

well. For example, if during detailed design it becomes apparent that a requirement cannot be met or can easily be exceeded, then the requirement may change. It is largely models (on the analysis/simulation side of Fig. 1) that can be used to play “what if . . .”. with a candidate design or operational concepts. The design engineer can use human performance models to determine whether requirements can or cannot be met.

In defining how human models could be used in detailed design validation, it is necessary to consider the kinds of information that would flow between the stages of the design process. Inherent in the idea of evaluating a design is the idea that it must be evaluated against some ‘target’, such as a performance standard or benchmark. Ideally, such standards and benchmarks should be created during concept validation (although the specific requirements could change during design iterations). While the whole system concept may include many benchmarks and data for system validation, those that are particularly relevant to the application of human models include:

- Lists of functions to be accomplished,
- Function allocations (to the extent that they have been allocated at any phase of design),
- Function level performance requirements (time, accuracy, risk),
- Manpower constraints,
- Personnel characteristic constraints, and
- Training constraints.

The *function lists* and *performance requirements* are not exclusively human functions at early phases of design when the allocation of functions among humans, hardware and software, (and even dynamic function allocation) is being considered. Indeed, much of the value of human performance models will be in assisting the tradeoffs of *function allocation* at all levels of design.

Also, in order to evaluate a design with respect to humans in the system, its constraints on the human subsystem must be known. These include *manpower constraints* (i.e. how many people), *personnel characteristic constraints* (i.e. what kind of people), and *training constraints* (i.e. how much time and resources will be available to provide the humans with the necessary skills for safe and successful system operation).

Without defining such constraints, it is impossible to reasonably validate or evaluate a design. For example, a system design might meet the manpower and personnel constraints, but it might take many years of training for the personnel to achieve the required level of performance. Therefore, all of these constraints must be defined before detailed engineering design begins.

Human models can have many uses in validation/evaluation of detailed system or component designs, such as:

- Predicting performance time,
- Predicting performance accuracy (e.g. error rates,% deviations),
- Predicting risk (e.g. probability of achieving function success),
- Predicting the satisfaction of anthropometric limitations,
- Predicting communication requirements/success,
- Predicting training requirements, and
- Integrating human performance predictions with other system performance models for integrated systems analysis.

Table 5 summarizes the models and modeling techniques that are applicable to each of these seven areas. It should be noted, however, that there are significant limitations in the capabilities of today's models in many or most of the boxes in Table 5. For example, in predicting and validating a design with respect to accuracy, there are many aspects of system accuracy that might need to be modeled, but only some can be effectively predicted using current human modeling technology.

Human Models to Support Design Synthesis

It is easy to envision the role that human performance models could play in supporting design analysis activities, as engineers and designers have been using models of hardware and software to do simulation-based system analysis for years. Design synthesis, however, includes what is often conceptualized as a creative component, and the application of human performance modeling techniques to support these types of activities is less intuitive.

The total-system and task-centered design views, as described above, can be used to help frame the kind of design tools that are needed to support the design process. For example, to support task-centered design, tools are needed to help identify team/organizational structures that match task requirements. An effort to develop such a tool is currently underway, supported by the Navy. Paley and colleagues (Paley, Levchuk, Serfaty & MacMillan, 1999) are developing a design tool (the Team Integrated Design Environment, or TIDE) that facilitates the application of a mission-based organizational design methodology. The first stage in following this methodology is to apply a multi-dimensional decomposition procedure to a particular mission, and determine the mission tasks and their interdependencies and relative sequencing. This mission structure can be represented in a series of mission task dependency graphs. This mission structure provides one of the two main inputs to the modeling tool. The other main input is a set of "organizational constraints", which include the

Table 5. Models Applicable to Analysis During Detailed Design.

Human Performance Model Class	Time	Accuracy	Risk	Anthropometric	Communication	Training	Integration
Closed form component models (e.g. Fitts' Law)	X	X					
Computational cognitive models	X	X			X	X	X
Critical decision models					X	X	X
Decision theory models			X				
Finite state models	X	X			X		
GOMS	X					X	
Group Training Model						X	
Human Reliability Analysis (HRA)		X	X				
Link Models of anthropometry movement				X			
Network models					X		
Optimal Control Models	X	X					
Perceptual Models	X	X					
Recognition Primed Decision Making (RPD)						X	
Signal Detection Theory		X					
Task hierarchies					X	X	X
Task Network	X	X			X		X
Training Taxonomy						X	
Workload	X	X					
Modeling Technique							
Knowledge Representation					X	X	
Decision-making & Reasoning			X		X	X	
Design/Analysis Techniques	X	X	X		X	X	X
Natural Language Processing Techniques					X		
General Purpose Simulation	X	X	X		X		X

resources and technologies available to support the human team members in accomplishing the mission. Given these two inputs, the modeling tool follows a three-part allocation algorithm to produce a team structure that is optimized to perform the mission. The tool produces an optimized structure by modeling organizational performance criteria as a multi-variable objective function, and then using advanced mathematical techniques to optimize that function. Of course, as with all design activities, this tool should be used in an iterative fashion, with multiple checks being performed on its output and the results of those checks fed back into the front end definition and analysis of the mission. While the software tool itself is still being developed, the underlying algorithms are already being successfully applied in several military domains.

To further support total-system and task-centered design, other tools are needed to analyze and determine when tasks and functions should be automated (replacing the human) versus when they should be supported by aiding-automation (keeping the person involved in the task). Issues such as determining effective task allocation, the impact of automation on human performance and the design of human-machine interfaces are addressed by practitioners in the field of Human Factors. In another Navy sponsored project, Eilbert and colleagues (Eilbert, Campbell, Santoro, Amerson & Cannon-Bowers, 1998) are currently using a cognitive computational modeling technology to capture some of the knowledge and reasoning capabilities of human factors engineers, in order to develop a design decision aid. (There is more discussion on the use of models to build decision aids, below.) The modeling framework chosen, COGNET, uses a combination of software components, including demons representing perceptual processes, a blackboard structure holding declarative knowledge, and GOMS-like rules capturing procedural knowledge, to simulate the opportunistic (or context sensitive) reasoning of a human expert in a limited domain. Once complete, the model will be used during design, not to replace human factors engineers, but rather to identify and draw to the attention of the system and design engineers the human factors issues, problems and analyses that need to be addressed. If this decision aid can increase the awareness of the existence and importance of human factors issues throughout the design of a complex system, then, hopefully, the human factors engineers will be given a larger role in supporting the design process.

Human Models and Human Experimentation

Traditionally, human issues in system design have been resolved, where possible, by experimentation with human subjects interacting with prototypes

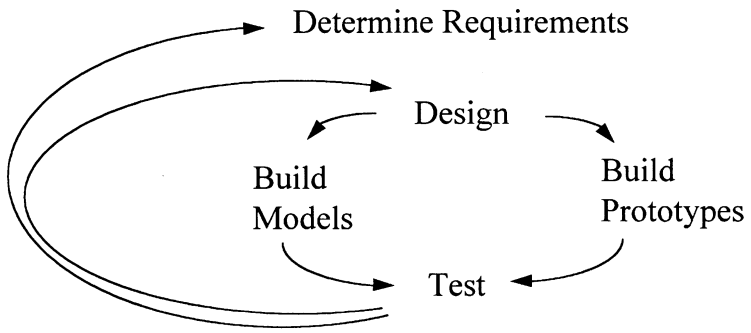

Fig. 2. Uses of Human Performance Models in Systems Validation.

or mockups or simulations of the system under design. The maturity and availability of human models will affect this practice. However, rather than reflecting an either/or relationship, these two approaches should be seen as having a complementary relationship. Figure 2 shows the roles that either human modeling or human-in-the-loop testing may play in validating either system requirements or design details.

The need for synergy between experimentation and modeling is perhaps even more important for human system components than elsewhere (e.g. hardware models) since the high degree of variability in human components of the system make them inherently less predictable. At the same time, the lower cost and logistical simplicity of models allows them to help focus the collection of empirical human performance data. A particularly important area in which this focusing could occur is risk mitigation. Models could be used effectively to focus the human-in-the-loop studies in the areas of design and operation that pose the greatest risk. Conversely, the human-in-the-loop studies could provide important calibration data for the human models, to enhance their predictive power and even to assess the validity of the models. This set of relationships is shown in Fig. 3.

Another linkage between model-based validation and human-in-the-loop testing is that the effectiveness of both techniques depends on the appropriate development and use of design basis scenarios in the validation process (a concept borrowed from the nuclear power industry). A design basis scenario is a scenario that taxes the system-under-design to the limits of its expected ability to be able to perform. Typically, several design basis scenarios are developed that push the system in different ways. To evaluate human/system performance and know when the human and system are performing acceptably,

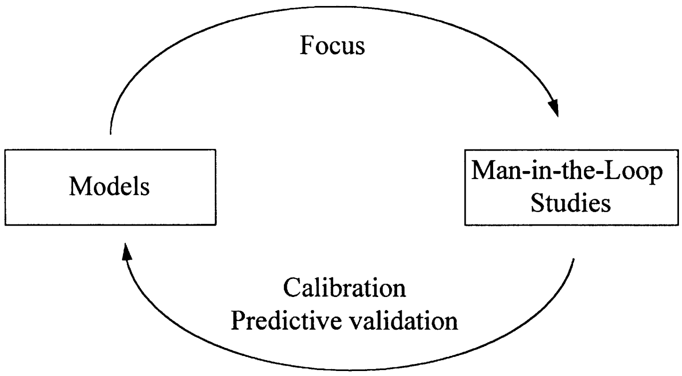

Fig. 3. Relationship between Human Performance Models and Man-in-the-Loop Studies.

there must be some human-focused design basis scenarios against which to test. Such scenarios will serve as the basis for model development *and* the conduct of human-in-the-loop experimentation.

General Issues Regarding Human Modeling Techniques to Support System Design

Regardless of whether human modeling techniques are being applied to support design synthesis or design analysis, there are three general issues that must be addressed: the level of design detail required for the application; the usability and interoperability of the modeling tool; and the maturity of the modeling technique. Each is further discussed below.

Many, if not all, variations of design process frameworks emphasize the iterative nature of design, in which design details slowly emerge over time. This can lead to a concern of when there is sufficient detail to begin applying human models. For example, is it necessary to know how many missiles a hypothetical person might be coordinating before that person/role can be modeled? Or how long it takes one missile to go from warm-up to detonate? The NYNEX corporation's experience in applying models to assess a new operator workstation design is illustrative here. Gray, John & Atwood (1993) were able to build the detailed GOMS models used in the evaluation of a new workstation only *after* the design was fully specified. Unfortunately, in the complex-systems acquisition process (particularly as practiced by DoD), once design details are specified, it is essentially too late to make changes if the modeling suggests that the design is sub-optimal. However, even though most of the historical use of human performance models in design has focused on human-computer interface issues, this does not have to be the case. For example, work is currently in process (Kirschenbaum, Gray & Ehret, 1997; Ehret, Kirschenbaum & Gray, 1998) to use models of command decision-making in submarine attack center design. What is useful in this case is modeling the decision-maker's information seeking strategies and behavior at a high level, not the details of how they actually get that information through the HCI of their workstations.

In fact, the human factors community argues strongly that human engineering must be done early and often during the design process, and that modeling can be useful before the implementation details of a design are determined. Fortunately, systems engineers are comfortable using high-level models of hardware and software (with the associated level of approximation in those models' outputs) early in the design process. But the questions of: (1) how one can build human performance models before detailed system design has begun, (2) how one can estimate the amount of imprecision in those models, and (3) the degree to which different modeling techniques are amenable to supporting high level models, are still somewhat open and debatable, making this an important area for further investigation.

It is also important to consider the potential impact (or lack thereof) of providing human performance models and modeling tools to systems engineers. It is unrealistic to expect to turn system engineers into human engineers simply by handing them a modeling tool. Ideally, a good tool will help in simple, straightforward cases, but in complex situations the support of a human factors engineer or human modeling specialist will likely be needed. In addition, the adoption of human-modeling tools will ultimately require use of a broader (human engineering) design process.

Interestingly enough, the degree of tool and model maturity needed before the tool/method becomes useful is, in fact, a variable. Returning to the NYNEX example, when Gray and John (1993) applied GOMS at NYNEX, the model was not developed as a part of a mature technology, yet the model nonetheless proved very useful (although it was highly sophisticated users – Psychology PhDs – who applied the technique). In general, a still-research-level model or tool can be made useful if the consumer organization has the commitment and willingness to put time and money into the process; typically this only happens if the problem involved is big and perceived as important. For the myriad of small, mundane problems, though, special resources will typically not be available, so highly mature and usable tools will be needed.

Finally, any human performance modeling tool must conform to the same rules as any other design tool. For example, if different human-model-based tools are required or recommended, then the tools should all be able to work from common inputs, so that the design at any particular level has to be captured only once. Designs captured at different levels of detail must also be traceable to one another. Modeling will have to prove more cost effective than prototyping combined with human-in-the-loop studies, and should enable more design options to be considered. Ultimately, the key to acceptance is the ability to assess operability and relate operability to cost. Model-based approaches and tools must be able to demonstrate that doing the additional analyses required by model-based approaches can significantly increase our ability to produce long run cost savings. And, of course, the tools themselves should be usable and well human-factored.

USING HUMAN MODELS IN SYSTEM OPERATION

The applications of models of human capabilities are not limited to the design and engineering of complex systems. It is also increasingly possible to embed cognitive and other human models in components of complex systems and use those models to support the operation of those systems. Another way to view this is as using the models as part of the implementation of the system, rather than as part of the design, as considered in the previous section. Following the approach of the design discussion above, the various functional ways in which human models could be used in system operation are first reviewed, and the current technology base from Table 1 is then mapped into the functional application areas.

The use of cognitive (or other human) models in system operation places a unique constraint on the development of a model that is not necessarily present when these models are being applied to system design. This is the constraint of *embeddability.* Applying a model in system operation implies that the model must exist in some form that is embeddable directly into the system itself. Purely analytical representations in prose or other non-executable formalisms may be used to inform the design process and may suggest specific features of the system being designed, but such representations cannot be actually embedded in the system itself.

The embeddability of a model implies some aspect of executability, but that executability can take many forms. The model can represent the internal processes of the human (for example, allowing the system to gain insights into human goals and intentions), or it could merely represent behavior (for example, manipulating a watchstation in a human-like manner). The model can deal with process, representing the internal steps in a decision or analysis process, or merely represent the outcomes via input-output relationships, but the constraint of embeddability still remains. Within this bounded definition of human model, ten functional applications to system operation, and the specific capability requirements associated with each application, can be defined.

(1) *Information access, retrieval, and integration*. In this application, the model is used to automate or replace human roles in gathering information and integrating it for use by human crew-members or other automated systems. Information access problems fall along a dimension that ranges from closed ended, in which the information to be retrieved is precisely defined from the problem conditions, to open-ended, in which defining the information to be sought and integrating partial pieces into a whole solution are part of the problem. Clearly, relatively simple human models can address closed-ended problems, while more open-ended problems would require more robust models that contain substantial knowledge and problem-solving capabilities.

(2) *Performance monitoring and assessment of human operators*. A model of normative or expected human performance can be used to monitor human crewmember's performance for adverse effects (e.g. of fatigue, extreme environmental condition, etc.) or other kinds of impairment, or it can be used to provide dynamic performance assessment of the human crewmembers for evaluation purposes, to identify training needs, etc. The relevant dimension for assessing the applicability of human models is the aspect of human performance that is being monitored or assessed. At one end is a purely behavioral monitoring/assessment (focusing on what the person is doing), and at the other end is a knowledge-oriented monitoring/ assessment (focusing on what the person knows). Here again, different models are needed for different points on this dimension, depending on their focus on cognition, behavior or both.

(3) *Real-time decision support*. Cognitive or decision models, typically encapsulating the strategy or knowledge of experts, can be used to aid the decision process or decision making of human crewmembers. This is the most commonly found application of embedded human models today. Based on the large amount of experience in this area, it does not appear that there are any inherent limits.

(4) *Associates*. A model or combination of models can be used to provide a digital assistant or one-on-one associate for a specific human crewmember (often in a senior or decision-making role). These associates can be used to off-load work in an on-demand basis, to support and simplify the

human-computer interaction with the watchstation, and/or to carry out other tasks when directed by the human being assisted. Operator associates is an area where substantial research has been invested to date (esp. DARPA associate systems programs), although much of the research has focused on technologies other than human models. The key constraint here is on the robustness of the associate, which must be able not only to perform many of the tasks assigned to the human, but also must be able to interact with the person in an intelligent and cooperative manner.

(5) *Embedded intelligent training*. Models of experts or instructors can be used to provide critiques of trainee actions, to define dynamically the correct or desired actions in training scenarios, and to specify the kinds of knowledge needed (or evidently lacking) for specific actions taken by trainees. The key distinction here is whether the training is knowledge based or performance based (or both). When only (behavioral) performance is being trained, then models that predict performance alone can be used. However when the goal is to diagnose the knowledge strengths/weaknesses of the trainee and focus instruction on those areas (rather than just on behavior), then models that represent knowledge and internal reasoning processes are needed.

(6) *Cooperation and collaboration support*. Models of team level processes and organizational work can be used to enhance, structure, and support collaboration and cooperation among teams of human crewmembers, for example by making sure that all required work gets done (i.e. nothing 'slips through the cracks') or that redundant work efforts are not being undertaken needlessly. The key dimension for assessing the applicability of human modeling technology here is the breadth of the cooperation. At one end is simple dyadic (one-on-one) collaboration, in which all work must be accomplished by that two-person team. At the other end are large organizations, in which there are many levels of restrictions concerning who can work with whom, information flows, etc. Nearer to the organizational end, models that deal with organization structures and processes (rather than individuals and roles) are needed, while nearer to the dyadic end, models that explicitly deal with individual responsibilities and characteristics are needed.

(7) *Dynamic role/function allocation*. Models of individual workload and organizational/team requirements can be used to re-allocate functions and roles (either among people on a team or between people and automation) to level workload and avoid performance bottlenecks due to overload of some (human or machine) parts of the system. Given that this definition is based on workload management, it is critical that models used be able

to assess the residual work capability of the human and machine agents involved. This assessment can occur either explicitly (e.g. through use of explicit workload models) or implicitly through a variety of other mechanisms. A secondary dimension of concern here is the ability to assess which agents are capable of performing specific functions that are candidates for re-allocation. Although lacking this facility, the dynamic reallocation process can be explicitly restricted to those functions that can be freely re-allocated to any agent within the team.

(8) *Task management*. Models of the flow of work (e.g. the flow of tasks or jobs needed to accomplish some key system-level goal) can be used to manage the efficient performance of the work under dynamic conditions, by reallocating functions, tasks, even whole human roles, as needed to get the job done. Task management differs from dynamic function allocation in that the latter, as defined here, is driven by individual workload issues (and is thus person-focused), while the former, as defined here, is driven by workflow issues, and is thus task focused. The considerations for task management are very similar to those for dynamic reallocation, with one exception. A task manager must still represent the workload and task-specific abilities of the agents (individuals or people) within its purview. However, because the task management process, as defined above, is task-flow oriented, models that can support this functional application must also have some explicit ability to represent and track the performance of tasks within an overall job-flow.

(9) *Task automation*. Models of human crewmembers that are sufficiently robust that they can perform tasks at approximately the same level (or perhaps even greater) than a human can be used as intelligent automation to replace humans for specific tasks or roles. There are relatively few restrictions on the use of models for task automation. Because task automation explicitly focuses on task performance (i.e. behavior), any model that can produce realistic and robust behavior is potentially applicable. In some cases, though, the ability to communicate about task performance may be required as well, and this would require use of models that can engage in explanation and inter-agent communication.

(10) *Knowledge management and transfer.* Finally, models can be used to act as acquirers and keepers of “corporate knowledge” and to disseminate or provide access to this corporate knowledge when needed. This emerging area is the least well-defined. The only clear constraint is that a model or modeling technique used for this function must have a clear and explicit representation of knowledge.

Given the requirements of the various applications delineated above, only certain modeling categories and techniques are appropriate for each application. The assessments are summarized in Tables 6 and 7. Table 6 focuses on specific human model types, while Table 7 focuses on more general modeling techniques.

CONCLUSIONS

This chapter has presented an assessment of the ways in which cognitive modeling and other related human modeling technologies could be used in the design, evaluation and operation of complex systems. Specifically, it has tried to delineate and characterize the short term opportunities for human model application in complex system engineering efforts, and to identify areas where additional research and development investment are needed and can lead to high-value applications in the future. Each of these goals is addressed below.

A commonly used metaphor is that of “low hanging fruit”, and this chapter tried to characterize the short-term application opportunities of human performance modeling technologies. Table 8 below summarizes these results. Using the taxonomy from Table 1, it reviews those techniques identified as ready for immediate or short-term application in each of the three areas of concern: system design, system operation, and concept and design evaluation. In addition, the current validation status of these techniques is also included.

Beyond the short-term applications, potential areas for high payoff research and development are also interesting. Many specific research needs and opportunities can be identified and classified into five general areas: cognition, knowledge management, team and organizational structure and processes, predictive models of training, and human-centered systems engineering. Each of these is further discussed below.

Research Need No. 1: Advanced Capability to Model Cognitive Processes

Research into cognition and development of models of cognition has led to many of the specific human models listed in Table 8. However, there are several areas where additional research support is needed before the technology can be applied to complex systems. These areas include:

(1) *Modeling human reading and understanding graphic-based displays*. The range of design options available today and in the future includes a dizzying array of interface technologies. Current cognitive models (even the cognitive modeling technologies that were developed for the purpose of human-computer interface evaluation) do not provide the ability to differentiate (in a predictive sense) human responses or performance across

Table 6. Models Applicable to System Operation Applications.

Model Type	Inf. Access, Retrieval, & Integration	Performance Assessment	Decision Support	Associates	Embedded Intelligent Training	Cooperation/ Collaboration Support	Dynamic Function Allocation	Task Management	Task Automation	Knowledge Management
Closed Form Component Models (E. G., Fitts' Law)			X							
Computational Cognitive Models	X	X	X	X	X	X*	X	X	X	X
Critical Decision Models			X							
Decision Theory Models			X							
Finite State Models	X	X	X	X	X		X	X	X	
GOMS			X							
Group Training Models										
Human Reliability Analysis (HRA)							X	X		
Network Models						X				
Optimal Control Models		X	X							
Perceptual Models			X							
Recognition Primed Decision-Making			X							
Signal Detection Theory			X							
Task Hierarchies										
Task Network	X	X	X	X	X	X	X	X	X	X
Training Taxonomy										
Workload		X	X							

* If expanded to include cooperative issues.

Table 6. Modeling Techniques Applicable to System Operation Applications.

Model Type	Inf. Access, Retrieval, & Integration	Performance Assessment	Decision Support	Associates	Embedded Intelligent Training	Cooperation/ Collaboration Support	Dynamic Function Allocation	Task Management	Task Automation	Knowledge Management
Blackboards	X		X	X	X	X*	X		X	X
Production Rules	X		X	X	X		X		X	X
Semantic Nets	X		X	X		X*				X
Bayesian Inference	X	X	X		X			X		
Fuzzy Logic	X		X		X			X		
Hybrid Logic			X					X		
Neural Nets		X	X				X	X	X	
Hidden Markov Models		X	X							
Case-Based Reasoning	X		X	X	X	X	X			X
Early Comparability Analysis										
Simulation-Based Tutoring & Authoring					X					
Simulation-Based Walk-Through										
Statistical Models (Descriptive)	X	X	X							
Statistical Models (Interpretive)										
Natural Language Processing	X	X	X	X	X	X	X	X	X	X
Monte-Carlo Methods			X		X					
Dynamical Models							X			

* By sharing the model across individuals.

Table 8. Near-Term Application Opportunities for Human Models.

Model Class	Design	Operations	Concept & Design Evaluation	Validation status (1)
Closed Form Component	X	X		face, construct, predictive
Computational Cognitive Models	X	X	X	face, construct, predictive
Critical Decision Models	X		X	face
Decision Theory Models		X		face, construct, predictive
Finite State Models		X		
GOMS	X		X	face, construct, predictive
Group Training Models				face
Human Reliability Analysis (HRA)			X	face, construct
Link Models Of Anthropometry & Mvmt			X	face, construct, predictive
Network Models	X			n/a
Optimal Control Models		X	X	face, construct, predictive
Perceptual Models			X	face, construct, predictive
Recognition Primed Decision-Making (RPD)				face, construct, predictive
Signal Detection Theory				face, construct, predictive
Task Hierarchies				face
Task Network	X	X	X	face, construct, predictive
Training Taxonomy	X			face
Workload			X	face, construct, predictive

these kinds of design alternatives, and thus cannot support the kinds of design decisions that must be made.

(2) *Extending computational cognitive architectures to provide collaborative support*. Current computational architectures lack the capabilities to model social/organizational cognition, metacognition, and discourse that are needed in truly collaborative work. These architectures need to be extended to include such capabilities so that models can be used in system

operation as workflow managers or as surrogate team-members, capable of adapting to changing team circumstances as human team members would.

(3) *Cognitive models of leadership*. Current cognitive models focus on task performance and decision making, yet leadership remains a key element of military, and indeed of all complex systems. The role of leadership in complex systems can not currently be modeled, and can not therefore be factored into system design or validation in a model-based way.

Research Need No. 2: Advanced Capability to Model Knowledge Management

Knowledge management is rapidly emerging as a key area in the engineering of manned systems, both commercial and military (e.g. Nonaka & Takeuchi, 1995). As an emerging area it is less well defined than cognitive modeling, but there are two areas where research needs can be identified:

(1) *Knowledge representation for engineering use*. Human factors engineers and systems engineers need access to the existing knowledge of complex system users to support future design, implementation, and evaluation activities. Techniques for representing specific kinds of human knowledge have been developed by cognitive researchers, but these methods were either oriented toward specific application frameworks (such as computational cognitive models) or for research and/or theoretical purposes. Research is needed to identify and implement methods and tools to make the existing (and future) knowledge representation schemes accessible to the system engineering process.

(2) *Knowledge growth/transfer.* Most knowledge representation methods and knowledge acquisition schemes treat knowledge as a static set of objects, to be captured and represented once. In reality, though, the knowledge of an individual or team is constantly changing. Some changes are the result of learning, and others are the result of environmental changes that cause some previously used elements of knowledge to become less useful (or even useless) and other elements to become more important. Organizational processes and system evolutions also constantly contribute to this process. Research is needed into means of capturing and representing these dynamic aspects of an individual or organizational knowledge base.

Research Need No. 3: Advanced Capability to Model Teams

Team/organizational research is also seriously needed in several areas. While much of the human modeling technology base developed with an individual-level focus (perhaps because of its roots in psychology), applications tend to revolve around designing and operating complex systems, for which key human behaviors frequently occur in team or organizational contexts. Three general areas can be singled out:

(1) *Team/organizational level modeling for cooperation/collaboration support*. Particularly in reduced manning systems, human teams must work together more effectively and efficiently. There will be little workload slack built into these systems, so the teams will have to collaborate more effectively, and in a much more complex and fluid environment than today, one in which roles, tasks, and even interfaces are dynamic and re-allocable. Surprisingly, there is relatively little useful research or (ideally) models concerning how teams collaborate (particularly at the cognitive level) that can be used to guide or even inform the design of collaborative environments or of automated tools such as task managers to support collaboration.
(2) *Team leadership*. Leadership is key to effective team performance, yet there are virtually no models of the leadership process, particularly in complex dynamic environments as envisioned, for example, in the future Navy ships. Models of team leadership are needed both to design and evaluate these systems.
(3) *Team level processes and training*. Cognitive research has produced useful models and architectures of individual level information processes and training. The team level analogs of these do not yet exist. Thus, the system designer has no clear reference points on the key limitations (e.g. analogous to short-term memory limits or cognitive biases) or important features in team level processes, particularly those of collaborative task performance and team training, making the design of work teams highly unstructured and subjective.

Research Need No. 4: Advanced Capability to Support Training Analyses Through Modeling

Several issues in the domain of training are also good candidates for future research. These concern not specific training methods or systems, but research into the training process itself so that training needs and costs could be better predicted during the system design and evaluation process. Specific modeling needs include:

(1) *Group training models*. Development of models of the processes by which groups or teams are trained, in terms of the underlying parameters and dependent variables such as time, cost, and effectiveness.

(2) *Training taxonomy for cost*. Research to map taxonomies of training techniques or training requirements (e.g. a skill taxonomy) into cost predictions, so that even at very early design stages, training options and cost could be estimated.

Research Need No. 5: Advanced Capability to Build Reusable, Interoperable Models

One final set of issues for future research concerns not human models per se, but the ways in which they mesh with the complex system engineering process. In many cases, these are issues associated with the transition of human modeling from a research to an engineering activity. They represent truly interdisciplinary applied research needs, and as such can easily fall into the cracks in the research funding process. However, they also represent critical problems that must be solved if the basic and exploratory research into model development is to be productively integrated into the system development process:

(1) *Model insertion*. Human models represent a new kind of tool for the systems engineering process. Even revolutionary systems such as the future Navy ships will be largely composed of legacy components, for which there are no human models, or at least for which no human models have been incorporated into the design/implementation/evaluation process. It is not clear how human models should be inserted into the life cycle of complex systems, nor what engineering and cultural issues associated with this insertion will arise.

(2) *Model integration*. Different human models must be able to be integrated with one another, and with the other models and tools used in the system development life cycle. This concern was echoed in the National Research Council report as well (Pew & Mavor, 1998). Research is needed to identify integration frameworks, to develop standards, and even to create integration support tools.

(3) *Reusable taxonomies/task description*. The lexicon of knowledge and human activities has never been standardized, or even cross-mapped among the many idiosyncratic frameworks used by different researchers or groups. Practitioners and appliers of this technology who are not researchers will need a taxonomy or reusable set of definitions and descriptions that they can use to identify and apply the appropriate human models. Research is needed either to standardize descriptions for concepts like task descriptions or taxonomies, or, where they do not exist, to create them. (One possible model for this process is the research to create a

unified medical terminology system that integrates and cross-maps the individual vocabularies among the many specialties of medicine.)

(4) *Tools to "bridge the gap" between system engineering functional decomposition tools and human-centered representations*. The tools of system engineering do not represent or decompose either structures or processes in the same way as human modeling methods such as task networks, computational cognitive models, etc. If the human models are to become part of the process, they will need to communicate with the existing system engineering decompositions, or make use of some higher-level framework that encompasses both needs. Without such a framework or translator, the cost of developing and managing dual decompositions for system and human modeling will prove too great a barrier.

(5) *Representation and modeling the evolvability/maintainability of systems & jobs*. Just as with knowledge, the jobs humans perform and the systems with which they perform them are constantly evolving, particularly during the post-deployment phase of the system life-cycle (which includes most of the life cycle, as well as most of the life-cycle costs). As the jobs and systems change, the design-phase analyses and models become increasingly irrelevant. Research is needed to find ways to incorporate this evolutionary aspect into the design phase models, and into the process of job/system evolution itself.

Of course, further follow-up is needed by design teams to make use of the available technology, and by the broader research organizations to fill the research gaps identified here. This is an exciting time for researchers and practitioners alike, as human modeling technology has now reached a state of maturity where it has proven itself able to become a contributing component to the design, evaluation and operation of complex systems.

NOTES

1. Many of the ideas contained in this paper originally arose from a three-day workshop on Cognitive Models in Complex Systems, sponsored by the Office of Naval Research and the U.S. Naval Air Warfare Center Training Systems Division. The authors express their gratitude to the sponsors of the workshop, and to the 45 researchers, engineers, and domain experts who contributed to the workshop. The authors also gratefully acknowledge the suggestions of Kevin Bracken on earlier drafts of this paper, as well as the contribution of Christine Volk in preparation of the manuscript.

2. The research into the "informant accuracy" problem is summarized in Bernard, Killworth, Kronenfeld and Sailer (1984).

3. Thus, the concepts of individual/network are maintained from the micro approach, but the use of organization-wide dynamics and the simplification of individual representations is retained from the macro approach.

REFERENCES

Anderson, J. R. (1976). *Language, Memory, and Thought*. Hillsdale, NJ: Erlbaum.

Anderson, J. R. (1993). *Rules of the Mind*. Hillsdale, NJ: Lawrence Erlbaum.

Baron, S., Kruser, D. S., & Huey, B. M. (1990). *Quantitative Modeling of Human Performance in Complex, Dynamic Systems*. Washington, D.C.: National Academy Press.

Bernard, H. R., Killwoth, P., Kronenfeld, D., & Sailer, L. (1984). The Problem of Informant Accuracy: The Validity of Retrospective Data. *Annual Review of Anthropology, 13*, 495–517. Ann Arbor: Annual Review Press.

Blanchard, B. S. & Fabrycky, W. J. (1998). *System Engineering and Analysis* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

Boff, K. R., Kaufman, L., & Thomas, J. P. (1986). *Handbook of Perception and Human Performance*. New York, NY: Wiley-Interscience.

Brown, J. S., & Slater-Hammel, A. T. (1949). Discrete Movements in the Horizontal Plane as a Function of their Length and Direction. *Journal of Experimental Psychology, 39*, 84–95.

Card, S., Moran, T., & Newell, A. (1983). *The Psychology of Human-Computer Interaction*. Hillsdale, NJ: Erlbaum.

Carnevale, Bost, R., Hamburger, T., Bush, T., & Malone, T. (June, 1998). Optimizing manning on DD 21. Paper presented at Warship ’98 International Symposium.

Carver, N., & Lesser, V. (1994). Evolution of blackboard control architectures. *Expert Systems with Applications, 7(1)*, 1–30.

Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior, 8*, 240–248.

Connelly, E. M. (1984). *Empirical investigation of aids for non-programming users in developing cost-effective requirements specifications*. Proceedings of the Human Factors Society Annual Meeting.

Czaja, S. J. (1997). Systems design and evaluation. In: G. Salvendy, (Ed.), *Handbook of Human Factors and Ergonomics* (pp.17–40). New York, NY: Wiley-Interscience.

Dodge, R., & Cline, T. S. (1901). The angle velocity of eye movements. *Psychology Review, 8*, 145–157.

Ehret, B. descriptive., Kirschenbaum, S. S., & Gray, W. D. (1998b). Contending with complexity: the development and use of scaled worlds as research tools. In: *Proceedings of the Human Factors and Ergonomics Society 42nd Annual Meeting*. Santa Monica, CA: Human Factors and Ergonomics Society.

Eilbert, J. L., Campbell, G. E., Santoro, T., Amerson, T. L., & Cannon-Bowers, J. A. (December, 1998). *The role of cognitive agents in the design of complex systems*. Paper presented at the 1998 Annual Interservice/Industry Training Systems and Education Conference. Orlando, FL.

Elkind, J. I., Card, S. K., Hochberg, J., & Huey, B. M. (1990). *Human Performance Models for Computer-Aided Engineering*. San Diego, CA: Academic Press, Inc.

Fitts, P. M. & Peterson, J. R. (1964). Information capacity of discrete motor responses. *Journal of Experimental Psychology, 67*, 103–112.

Forrester, Jay. (1971). *Principles of Systems*. Cambridge, Mass.: MIT Press.

Giffen, H. C., III, VADM (1997). *Smart Ship Project Assessment Report*. http://www.dt.navy.mil/smartship/assess0997.html.

Gray, W. D., John, B. E., and Atwood, M. E. (1993). Project Ernestine: Validating a GOMS analysis for predicting and explaining real-world task performance. *Human-Computer Interaction*, *8*, 237–309.

Green, D. M., & Swets, J. A. (1966). *Signal Detectability and Psychophysics*. New York, NY: Wiley.

Hutchins, C. W. (1974). *A computer-aided function allocation and evaluation system (CAFES)*. Proceedings of the Human Factors Society Annual Meeting.

IEEE Standard 1220–1998, IEEE standard for applications and management of the systems engineering process (1999). New York: Institute of Electrical and Electronic Engineers.

Kalman, R., Falb, P., & Arbib, M. (1969). *Topics in Mathematical Systems Theory.* New York: McGraw Hill.

Kieras, D. E. Meyer, D. E. (1997). An overview of the EPIC architecture for cognition and performance with application to human-computer interaction. *Human-Computer Interaction*, *12*(4), 391–438.

Kirschenbaum, S. S., Gray, W. D., & Ehret, B. D. (1997). *Subgoaling and subschemas for submariners: Cognitive models of situation assessment* (Technical Report NUWC-NPT TR 10,764–1). Newport, RI: Naval Undersea Warfare Center Division Newport.

.Klein, G. (1989). Recognition-primed decisions. In: W. Rouse (Ed.), *Advances in man-machine systems research*, *5*, 47–92. Greenwich, CT: JAI Press, Inc.

Klein, G. A., Calderwood, R., & MacGregor, D. (1989). Critical decision method for eliciting knowledge. *IEEE Transactions on Systems, Man, and Cybernetics*, *19*(3), 462–472.

Laird, J., Coulter, K., Jones, R., Kenney, P., Koss, F., & Nielsen, P. (1997). *Review of Soar/FWA Participation in STOW-97*. Published electronically at http://ai.eecs.umich.edu/ifor/stow-review.html. November 7, 1997.

Laird, J. E., Newell, A., & Rosenbloom, P. S. (1987). Soar: An architecture for general intelligence. *Artificial Intelligence*, *33*(1), 1–64.

Lane, N., Strieb, M., Glenn, F., & Wherry, R. (1981) The Human Operator Simulator: An overview. In: J. Moraal & K.-F. Kraiss (Eds), *Manned Systems Design: Methods, Equipment, and Applications*, New York: Plenum Press.

Laughery, K. R. (1998). Modeling human performance during system design. In: E. Salas (Ed.), *Human/Technology Interaction in Complex Systems*. JAI Press.

Laughery, K. R. & Corker, K. (1997). Computer modeling and simulation. In: G. Salvendy's (Ed.), *Handbook of Human Factors and Ergonomics*. New York, NY: John Wiley & Sons

McMillan, G. R., Beevis, D., Salas, E., Strub, M. H., Sutton, R., & Van Breda, L. (Eds) (1989). *Applications of Human Performance Models to System Design*. New York, NY: Plenum Press.

McNeese, M. D., Warren, R. & Woodson, B. K. (1985). *Cockpit automation technology: a further look*. Proceedings of the Human Factors Society Annual Meeting.

Meadows, D., Meadows, D., Randers, J., & Behrens, W. W., III. (1972). *The Limits to Growth*. New York: Universe Books.

Moraal, J., & Kraiss, K. F. (Eds). (1981). *Manned Systems Design: Methods, Equipment, and Applications*. New York: Plenum Press.

Nonaka, I., & Takeuchi, H. (1995). *The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation*. London: Oxford University Press.

Paley, M. J., Levchuk, Y. N., Serfaty, D., & MacMillan, J. (1999). *Designing optimal organizational structures for combat information centers in the next generation of Navy* *ships*. Paper submitted to 1999 Command and Control Research and Technology Symposium.

Pew, R. W., & Mavor, A. S. (Eds) (1998). *Modeling Human and Organizational Behavior: Application to Military Situations*. Washington, DC: National Academy Press.

Risser, D. T. & Berger, P. (1984). *Army HARDMAN: Its origin, evaluation, implementation to date*. Proceedings of the Human Factors Society Annual Meeting.

Rumelhart, D. E., McClelland, J. L., & The PDP Research Group (1986). *Parallel Distributed Processing: Explorations in the Microstructure of Cognition* (2 volumes). Cambridge, MA: MIT Press.

Siegel, A. I. & Wolf, J. J. (1962). A model for digital simulation of two-operator man-machine systems. *Ergonomics*, *5*, 557–572.

Siegel, A. I. & Wolf, J. J. (1969). *Man-Machine Simulation Models*. New York, NY: Wiley-Interscience.

Swezey, R. W. & Llaneras, R. E. (1997). Models in Training and Instruction. In: G. Salvendy (Ed.), *Handbook of Human Factors and Ergonomics*. New York, NY: Wiley-Interscience, 514–577.

Von Neumann, J., & Morgenstern, O. (1947). *Theory of Games and Economic Behavior.* Princeton University Press, Princeton, NJ.

White, H., Boorman, S., & Breiger, R. Social structure from multiple networks. I. Blockmodels of roles and positions. *American Journal of Sociology, 81*, 730–780.

Wickens, C. D. (1980). The structure of attentional resources. In: R. Nickerson & R. Pew (Eds), *Attention and Performance VIII*. Hillsdale, NJ: Erlbaum.

Wortman, C. D., Duket, S. D., Seifert, D. J., Hann, R. I., & Chubb, A. P. (1978). *Simulation using Saint: A user-oriented instruction manual*. Wright-Patterson AFB, OH: Aerospace Medical Research Laboratory (AMRL-TR-77–61).

Zachary, W. (1977). An information-flow model for conflict and fission in small groups. *Journal of Anthropological Research, 33* (Winter), 452–73.

Zachary, W., Bilazarian, P., Burns, J., & Cannon-Bowers, J. (1997). Advanced Embedded Training Concepts for Shipboard Systems. In: *Proceedings of 19th Annual Interservice/Industry Training Systems and Education Conference*.

Zachary, W., Le Mentec, J. C., & Ryder, J. (1996). Interface agents in complex systems. In: C. Ntuen & E. H. Park (Eds), *Human Interaction With Complex Systems: Conceptual Principles and Design Practice*. Norwell, MA: Kluwer Academic Publishers.

Zachary, W., Cannon-Bowers, J., Burns, J., Bilazarian, P., & Krecker, D. (1998). An Advanced Embedded Training System (AETS) for tactical team training. In: *Proceedings of ITS '98*.

Zachary, W., Ryder, J., Ross, L., & Weiland, M. (1992). Intelligent Human-Computer Interaction in Real Time, Multi-tasking Process Control and Monitoring Systems. In: M. Helander & M. Nagamachi (Eds), *Human Factors in Design for Manufacturability* (pp. 377–402). New York: Taylor and Francis.

Zadeh, L. A. (1965). Fuzzy sets. *Information and Control, 8*, 338–353.